BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 3: e36125
doi: 10.3897/biss.3.36125

OPEN ACCESS

Conference Abstract

The SILVA Database Project: An ELIXIR core data resource for high-quality ribosomal RNA sequences

Frank Oliver Glöckner ‡

‡ Alfred-Wegener-Institut, Helmholtz Zentrum für Polar- und Meeresforschung and Jacobs University, Bremen, Germany

Corresponding author: Frank Oliver Glöckner (frank.oliver.gloeckner@awi.de)

Received: 12 May 2019 | Published: 21 Jun 2019

Citation: Glöckner F (2019) The SILVA Database Project: An ELIXIR core data resource for high-quality ribosomal RNA sequences. Biodiversity Information Science and Standards 3: e36125. https://doi.org/10.3897/biss.3.36125

Abstract

Ribosomal DNA (rDNA) has become the primary target molecule for phylogenetic reconstruction and the cultivation-independent detection and quantification of microorganisms (barcoding). With the advent of high-throughput sequencing technologies (Next Generation Sequencing (NGS), PCR-based amplicon sequencing of rDNA fragments for diversity screening is now a routine technology, at least in environmental sciences. The resulting exponential increase of publicly available rDNA sequences demands specialized reference databases (Fig. 1).

Figure 1.

SILVA Logo

© Glöckner F. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

SILVA (from Latin *silva*, meaning forest) is designed to provide a comprehensive web resource for up-to-date, quality-controlled databases of aligned rDNA sequences from the Bacteria, Archaea and Eukaryota domains and the corresponding online services (Glöckner et al. 2017, Quast et al. 2012).

The current SILVA database (release 132) contains 6,073,181 small subunit and 907,382 large subunit rRNA gene sequences. All sequences are checked for anomalies, carry a rich set of sequence-associated contextual information, multiple taxonomic classifications (EMBL-EBI/ENA, RDP and GTDB) and the latest validly-described nomenclature. SILVA maintains manually curated reference alignments of 75,000 ribosomal RNA genes, both 16S/18S (small subunit, SSU) and 23S/28S (large subunit, LSU). With every full release, a manually curated guide tree is provided that contains the latest taxonomy and nomenclature based on multiple references.

SILVA is the only rDNA database project worldwide where special emphasis is given to the consistent naming of clades of uncultivated (environmental) sequences where no validly-described cultivated representative is available (Yilmaz et al. 2014). SILVA incorporates other unique features, including a comprehensive 23S/28S database of aligned rDNA sequences and alignments that contain Eukaryota sequences.

SILVA is an active partner of RNACentral. RNAcentral is a public resource that offers integrated access to a comprehensive and up-to-date set of non-coding RNA sequences provided by a collaborating group of Expert Databases. The SILVA team is a member of the Bergey's Board of Trustees which provides the authoritative taxonomy for *Bacteria* and *Archaea* as well as the Protist Reference Taxonomy Project UniEuk funded by the Gordon and Betty Moore Foundation to create a unified taxonomic framework

In 2018 SILVA became an ELIXIR Core Data Resource. ELIXIR Core Data Resources are a set of European data resources of fundamental importance to the wider life-science community and the long-term preservation of biological data.

To facilitate classification tasks for high-throughput rDNA data the SILVAngs has been implemented and released in 2013. SILVAngs is a data analysis service for rDNA reads from high-throughput sequencing (NGS) approaches based on an automatic software pipeline. It uses the SILVA rDNA databases, taxonomies, and alignments as a reference. It facilitates the classification of rDNA reads and provides a wealth of results (tables, graphs and sequence files) for download. SILVAngs serves several thousands of registered users, which processes thousands of projects per year.

The application spectrum of the SILVA databases ranges from environmental sciences, microbiology, agriculture, biochemistry, biotechnology to medicine in academia and industry.

Keywords

biodiversity, ribosomal RNA gene, rDNA, alignments, taxonomy, phylogeny, databases, probes, primers, amplicon (NGS) analysis

Presenting author

Frank Oliver Glöckner

Presented at

Biodiversity_Next 2019

Acknowledgements

We would like to thank all SILVA and SILVAngs users for their support and feedback on the SILVA databases and tools. Your immediate critical evaluation is a great motivation for us to improve our databases, tools and services. Special thanks goes to the whole de.NBI network where foreigners turned into friends inspired to serve the scientific community at its best.

References

- Glöckner FO, Yilmaz P, Quast C, Gerken J, Beccati A, Ciuprina A, Bruns G, Yarza P, Peplies J, Westram R, Ludwig W (2017) 25 years of serving the community with ribosomal RNA gene reference databases and tools. Journal of Biotechnology 261: 169-176. https://doi.org/10.1016/j.jbiotec.2017.06.1198
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2012) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research 41: D590-D596. https://doi.org/10.1093/nar/gks1219
- Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, Schweer T, Peplies J, Ludwig W, Glöckner FO (2014) The SILVA and “All-species Living Tree Project (LTP)” taxonomic frameworks. Nucleic Acids Research 42: D643-D648. https://doi.org/10.1093/nar/gkt1209